Examining the Effect of an Interdisciplinary Literacy Program on Student Learning

Jia Wang¹ · Joan L. Herman¹ · Seth Leon¹ · Scott Epstein¹

Accepted: 31 May 2021 / Published online: 14 June 2021
© The Author(s) 2021

Abstract

The literacy design collaborative (LDC) was created to support teachers in implementing college and career readiness standards. Teachers work collaboratively with coaches to further develop their expertise and design standards-driven, literacy-rich writing assignments within their existing curriculum across content areas. The current paper reports on early student academic outcome results from a multi-year mixed methods study of the implementation and effect of LDC using a quasi-experimental design, as implemented in one large urban school district. We found LDC had statistically significant effects on Cohort 2 middle school students' English language arts (ELA) performance, an effect size of ($d=0.15$) for students receiving LDC instruction in the three core content areas LDC focuses on: ELA, science, and social studies/history.

Keywords Student learning · Quasi-experimental design · Literacy intervention · Interdisciplinary collaboration · College and career readiness standards implementation

Introduction

The literacy design collaborative (LDC) was created to support teachers across content areas in their classroom implementation of college and career readiness standards in 2011. Teachers work collaboratively with their peers and coaches to further develop their expertise in designing and implementing curriculum centered around standards-driven writing assignments. The LDC intervention consists of four components: a coach-supported Professional Learning Community (PLC); asynchronous

✉ Jia Wang
jwang@cresst.ucla.edu

¹ National Center for Research on Evaluation, Standards, and Student Testing (CRESST), Graduate School of Education and Information Science (GSE&IS), University of California Los Angeles, 300 Charles E. Young Dr. North, GSEIS Bldg. Rm. 307, Los Angeles, CA 90095-1522, USA

support from coaches; implementation activities; and leadership support at different levels. Since its inception, LDC has been widely used by individual teachers, schools, and districts across the country and adopted as a statewide strategy for Common Core implementation in Kentucky, Colorado, Louisiana, and Georgia.

Prior research on LDC has provided early evidence that teachers find LDC useful, that LDC improves teachers' skills, and that LDC has a positive effect on student learning, at least in certain contexts (e.g. Levin & Poglinco, 2013; Herman et al., 2016). The LDC model is currently implemented in two large urban school districts as part of Investing in Innovation Fund (i3) grant. The current paper reports on early student academic outcome results from this i3-funded multi-year mixed methods study of LDC's implementation and impact on student learning using a quasi-experimental design (QED), as implemented in one large urban school district.

Overview of LDC Intervention

College and career readiness standards (CCRS) have raised the bar on what students are expected to learn, placing higher demands on teacher practice and capacity (Darling-Hammond et al., 2014). One of the primary goals of LDC is to provide specific guidance to help teachers across the curriculum to integrate English language arts CCRS effectively into their classroom instruction. LDC does so by providing task templates, training, and other supports that enable teacher to create and implement challenging content- and literacy rich units that can be seamlessly embedded in classroom curriculum. The templates assist teachers in creating extended units that culminate in a content-oriented, evidence-based writing assignment. The units present clear expectations for reading and writing, as well as an instructional ladder designed to help students build the literacy and content understanding skills they need for successful completion of the unit.

LDC thus supports one of the key shifts in CCRS for ELA: that students engage with complex, content-rich texts across multiple disciplines (Chadwick, 2015). Integration of literacy within content area programs has shown increased student learning in content knowledge, vocabulary, writing, and reading comprehension (Duesbery et al., 2011; Goldschmidt, 2015; Reisman, 2012). By explicitly encouraging and assisting teachers in other content areas to integrate their content standards with literacy standards (Carter et al., 2007; De La Paz et al., 2016; Draper, 2008; Goldschmidt, 2015; Klein & Kirkpatrick, 2010; Monte-Sano & De La Paz, 2012), LDC supports the teaching of reading and writing for all teachers, considering literacy development a shared responsibility throughout the school (Shanahan & Shanahan, 2008).

One of the key instructional design principles of LDC is backward mapping (Graff, 2011; Wiggins & McTighe, 2005). After identifying a high-quality writing assignment based on content standards and CCRS, teachers separate the task into the specific content skills and literacy skills that students need in order to succeed on the task. Teachers formulate a unit by keeping in mind the learning goal, the necessary steps to achieve the goal, and the vision of high quality performance while planning and enacting lessons (Ball, 2000). The LDC design process serves

as the overarching framework of practice and teachers have opportunities to share representations, practice decompositions, and view approximations of practice while learning to use specific scaffolds and approaches to instruction (Grossman et al., 2009).

LDC is providing selected schools and teachers in two urban districts with training and support to implement the LDC model through a federal i3 validation grant. Participating district participants, mainly classroom teachers, learn how to utilize the template tools and backward design process described above by participating in a PLC in their school. Each PLC is supported by a coach who joins the team digitally every other week and provides feedback within the LDC online platform, called CoreTools. Teachers are expected to adapt/create and teach at least two LDC modules over the course of each school year. In addition to the module building platform, teachers also have access to an extensive library of exemplars, tools, templates, and resources in CoreTools. Additionally, school and teacher leaders are engaged in a process of goal setting and reflection related to the LDC work.

This study is of vital importance as it could provide evidence on a promising program designed to improve students' literacy skills. Overwhelming data demonstrate that American students lag in essential literacy achievement. For example, American 15-year olds were more likely to score at below level 2 on the PISA than students in 33 of 65 participating nations. LDC offers an integrated approach improving literacy instruction by increasing the rigor of content, the skills and knowledge of teachers, and the active engagement of students in learning. Providing evidence on the program's efficacy can benefit districts, schools, and teachers searching for comprehensive approaches to raise the quality of learning in their schools.

Study Methodology

The current paper reports on early student academic outcome results from a multi-year mixed methods quasi-experimental study of the implementation and impact of LDC. The full evaluation study is a comprehensive mixed-methods evaluation to understand the impact of LDC on students, as well as to document LDC's impact on teacher skills and practices, and draws on data from two cohorts of schools from two large urban school districts. The evaluation study measures teacher implementation and skill improvement with teacher surveys, analytic data from LDC's online CoreTools module building platform, and artifact analysis of the modules LDC teachers developed and implemented in the classrooms. At the time of this paper, we are in the process of receiving 2018–2019 student data from one urban school district and are expecting to receiving the same data from the second urban school district in January 2020.

In this paper, we will focus on the effect of LDC intervention on the academic performance of participating students as measured by the state assessment in ELA in 2017–2018. Specifically, we employed a quasi-experimental design to examine the effect of LDC on the Smarter Balanced Assessment Consortium (Smarter Balanced) ELA assessment scores of students in elementary and middle schools participating in the LDC program. We used a two-step matching process to establish the

baseline equivalence between treatment students and a reduced pool of comparison students and teachers at schools with similar characteristics.

A three-level analysis model is employed to examine the effectiveness of LDC in increasing student learning for Cohort 1 schools with 2 years of LDC implementation and Cohort 2 schools with 1 year of implementation. Our multilevel models incorporate demographic and achievement variables used in the matching design as covariates, making the findings “double robust” in that characteristics are controlled for in both matching and outcomes analysis stages.

Sample Recruitment and Sample Description

This paper examines LDC as implemented in a large urban district using 2017–2018 student outcome data. The sample consists of 14 schools which began their LDC intervention during 2016–2017 (Cohort 1), and 31 schools which commenced at the beginning of the 2017–2018 school year (Cohort 2). The LDC program team recruited participating schools, which then constituted PLCs made up of teacher. We worked with the LDC program team to recruit the PLC teachers for the research study. We began the teacher recruitment at the beginning of the school year with authorized staff attending LDC Launch Days. We continued recruitment and consent through in person contact, video conference, and email. All participating teachers, teacher leaders, and administrators were compensated with a \$50 gift card after completing the study survey.

Among the total of 350 LDC participants in the school year 2017–2018, 286 were classroom teachers and the remaining participants were either teacher leaders or administrators who supported the intervention work at the school sites but were not classroom teachers and therefore were not linked to students via courses. As the school district required individually signed consent forms before identifying these LDC teachers in the district data, our outcome analysis only analyzed data on teachers who consented to participate in the study. A total of 271 classroom teachers consented to participate in the study, for a consent rate of 94.8%. Please see Table 1 for the breakdown of participating classroom teachers by their school Cohort, years of LDC implementation, and their school level.

To conduct the student outcome analysis, the sample was further restricted by the need for student achievement data for both the outcome year (2017–2018) and the baseline year (2015–2016 for Cohort 1 schools/students and 2016–2017 for Cohort 2 schools/students). Additionally, teachers teaching either in high school or the primary elementary grades (K–3) were not included in the student outcome analyses due to the lack of baseline achievement scores, as required for conducting quasi-experimental studies. Middle school teachers who did not teach a core ELA, science, or social studies/history class were also excluded from the analyses because LDC is expecting to make a meaningful impact in these three identified core subjects.

We analyze the elementary and middle school data separately because the LDC intervention was originally developed and validated for middle schools and implementation at the elementary school represents an expansion of the scope of LDC. In addition, given LDC’s theory of action that the impact of LDC will be felt after at

Table 1 Number of classroom teachers in 2017–2018 by cohort, implementation years, and school level

Cohort/years of implementation/school level	Participated in LDC in 2017–2018	Consented to participate in study	Consented teachers with eligible students
<i>Cohort 1 Schools Starting in 2016–2017</i>			
Elem Teachers Implemented LDC in 2017–2018 only	20	20	1
Elem Teachers Implemented LDC in 2016–2017 and 2017–2018	29	29	3
Mid Sch Teachers Implemented LDC in 2017–2018 only	6	6	4
Mid Sch Teachers Implemented LDC in 2016–2017 and 2017–2018	22	22	22
High Sch Teachers Implemented LDC in 2017–2018 only	7	6	0
High Sch Teachers Implemented LDC in 2016–2017 and 2017–2018	4	4	0
<i>Cohort 2 Schools Starting in 2017–2018</i>			
Elem Teachers Implemented LDC in 2017–2018	138	129	85
Mid Sch Teachers Implemented LDC in 2017–2018	54	50	43
High Sch Teachers Implemented LDC in 2017–2018	6	5	0
Total	286	271	158

least 2 years of teacher participation, we conducted analysis for teachers with 1 or 2 years of implementation experiences separately.

As a result, quasi-experimental analyses were conducted for three groups of LDC intervention teachers with sufficient sample size: Cohort 1 returning middle school teachers ($n=22$), Cohort 2 new elementary teachers ($n=85$), and Cohort 2 new middle school teachers ($n=43$). At the time of Spring 2018 testing, returning teachers would have implemented LDC in their classes for 2 years and new teachers would have implemented LDC for 1 year. The last column in Table 1 reports the number of LDC teachers who were teaching students in grades 4–8 and therefore were eligible for our quasi-experimental design analyses.

Data and Matching Process

We requested and received the administrative data on teachers participating in the LDC PLCs, and their students (including demographics and state assessment scores) from the participating district annually. Student-level variables utilized in the outcome analysis included race/ethnicity, gender, poverty status, special education status, English language proficiency, gifted status, grade, baseline achievement in mathematics and ELA, and outcome year achievement in ELA on state assessments. Teacher-level indicators obtained and utilized included years of teaching experience and teacher attendance. We also requested and received roster files that establish a link between teachers and students via specific courses.

We then used a two-step matching process to identify comparison students and teachers at schools with similar characteristics to the schools in the intervention sample. To accomplish this, we first identified the four or five most similar comparison schools for each intervention school based on a Euclidian distance measure, by using the nearest neighbor analysis option in SPSS 24.0 (see Fix & Hodges, 1951; Wang et al., 2007). The variables used in this process were the percentage of students eligible for free or reduced price lunch, the percentage of African American students, mean baseline student achievement in ELA, mean baseline student achievement in mathematics, the average attendance rate of teachers, the percentage of teachers with three or fewer years of teaching experience, and the school grade span where feasible.

The student-level matching technique we employed was coarsened exact matching (CEM) (Iacus et al., 2011). CEM is a flexible matching approach with many favorable properties and allows the researcher to specify the precise conditions under which students are matched. For categorical variables, such as race/ethnicity or free or reduced price lunch status, this can entail exact matching, while for continuous measures, such as baseline individual student achievement and aggregate class level achievement, cut-points for matching can be specified. With this approach we were able to set precise cut-points on the most important baseline indicators, such as baseline academic achievement, to ensure that where possible every treatment student was matched with a suitable comparison. Student matching variables we used in CEM included Hispanic, Black, poverty status, female, English language

proficiency (English language learner), special education status, gifted status, mean baseline achievement in mathematics and ELA, and grade level.

During matching we also included a few variables capturing information on the teachers and peers to whom students were exposed, given research on the effects of school context on student performance (e.g. Zhu et al., 2012). These variables included mean baseline ELA achievement of the student's peers in his/her core content classes, and the average years of teaching experience of the student's core content teachers. For each of the three groups of students for whom outcome analyses were conducted, we retained between 90 and 93% of the eligible students. The purpose of this comprehensive matching process was to ensure the resulting sample resembles the type of sample one would expect to obtain through random assignment.

Calculating the LDC Dosage Variable at the Student Level

In both elementary and middle school datasets, students were linked to their teachers through course enrollment data and the LDC dosage was assigned as a proportion of a unity. For elementary students, the teacher to whom a student was assigned for each of the three marking periods was used. If a student was exposed to one LDC teacher across all three periods, the student's LDC exposure would add to a unity (1). If a student was exposed to one LDC teacher for two periods and one non-LDC for one period, the student's LDC exposure would be 0.67.

Middle school students are assigned to classes (associated with teachers) by content area for each of two semesters. Core content areas for LDC participation were ELA, social studies/history, and science. Student assignments in LDC core areas were used to establish dosage weights. The weighting in middle school was always distributed as a proportion of the total semesters across the three content areas. Therefore, if a student accumulated one science unit (one semester), two social studies/history units (two semesters), and two ELA units (two semesters), the base number of units would be five. Using that scenario, the science teacher would contribute one fifth (0.20) of the overall core curriculum exposure with the social studies and science teachers contributing two fifths (0.40) each, again resulting in the student's exposure adding to a unity.

For this study, we modeled the LDC dosage variable at the student level in two ways. The first dosage-dependent approach takes into account the students' level of exposure to the LDC intervention teachers. In this approach, the treatment was structured as a continuous response variable, coded as zero for comparison students and coded as a positive value for treated students, albeit never exceeding one. The positive value assigned to treated students in the dosage-dependent approach was simply the sum of the intervention teacher weights linked to the treated student. The second approach was dosage independent and classified any student exposed to any LDC intervention teacher via at least one course as a treated LDC student. In this approach the treatment variable was dichotomous: coded as one for treated students and zero for comparison students.

Analysis Approach

We employed a quasi-experimental design to examine the effect of LDC on student test scores on the Smarter Balanced Assessment Consortium (Smarter Balanced) ELA assessment in the participating LDC elementary and middle schools in 2017–2018. We conducted separate analysis for elementary and middle schools, and for teachers with 1 or 2 years of experience in implementing LDC in their classrooms. As students—especially middle school students—were likely to have connections to multiple teachers in the available time period prior to each testing outcome (students at the elementary school level were also sometimes exposed to multiple teachers, but to a lesser extent), the LDC effects were estimated using multiple membership multiple classification (MMMC; Browne et al., 2001). MMC is an extension of the standard multilevel modeling framework and the most appropriate modeling approach as students can be exposed to LDC-infused instruction via multiple content area teachers at the secondary level. MMC has the flexibility to account for the full range of teacher/student exposures that occur in secondary settings.

A three-level MMC model was used to estimate the effects of the LDC intervention on student learning. The general specification is shown in the following equation using similar notation proposed by Browne et al., (2001, Eq. 6) and applied in Tranmer et al. (2014, Eq. 3).

The three lines together make up the equation and is part of the text.

$$y_i = x_i' \beta + u_{School(i)}^{(3)} \sum_{j \in Teacher(i)} w_{ij} u_j^{(2)} + e_i$$

$$i = 1, \dots, n \text{ Teacher}(i) \subset (1, \dots, J)$$

$$u_{School(i)}^{(3)} \sim N(0, \sigma_{u(2)}^2), u_j^{(2)} \sim N(0, \sigma_{u(2)}^2), e_i \sim N(0, \sigma_e^2)$$

In this model y_i is the student achievement score response, X_i is a vector of the fixed covariates, and β is the vector of the corresponding fixed effects. $School(i)$ is the school which student i attends, thus the term $u_{School(i)}^{(3)}$ represents the random effects for that level of classification. Within the term $\sum_{j \in Teacher(i)} w_{ij} u_j^{(2)}$, $u_j^{(2)}$ is the set of j random effects for the teachers included in the selected dataset, and w_{ij} is the weight which sums to 1 for each student applied in proportion to the instruction time assigned with each teacher.

Student assessment outcomes are measured in standardized units based on the raw scores in the Smarter Balanced ELA assessment. The standardized scores are based on the district mean and standard deviation for each grade level, allowing us to compare scores across grades more easily and compatibly.

As presented in Table 1, the MMC analysis was only conducted for three groups of teachers (and their students) due to sample size restrictions. The three specific groups of teachers we analyzed were: Cohort 1 middle school teachers in their

second year of LDC implementation, Cohort 2 elementary teachers in their first year of implementing LDC, and Cohort 2 middle school teachers in their first year of implementation.

Analysis Results

Tables 2, 3 and 4 display the results of QED analyses estimating the effect of LDC on students' ELA state assessment scores in 2017–2018. In Table 2, we present results of both the dosage dependent and dosage independent models on cohort 1 middle school students' ELA performance in 2017–2018, after being taught by teachers with 2 years of LDC experience. As can be seen, the LDC effect on student outcomes is in the positive direction, but not statistically significant for either model. In other words, neither analysis met statistical criteria for concluding that students taught by LDC teachers performed better on the ELA test than did their matched peers in the comparison group.

Statistically significant effects of the covariates on student performance were similar under the two models and were in the expected directions. Baseline ELA performance was the strongest predictor and baseline mathematics performance also helped explain the outcome. In addition to baseline achievement, three demographic variables helped predict performance: English language learners performed at lower levels than English Only and Reclassified Fluent English Proficient students, females performed at significantly higher levels than males, and students enrolled in

Table 2 Cohort 1 returning LDC middle school teacher effect estimates on 2017–2018 smarter balanced ELA performance, dosage-dependent and dosage-independent models

Variables	Dosage-dependent model coefficient (<i>SE</i>)	Dosage-independent model coefficient (<i>SE</i>)
<i>Level 2—LDC teacher treatment</i>	0.118 (0.116)	0.050 (0.083)
<i>Level 1—student characteristics</i>		
Hispanic	−0.216 (0.127)	−0.215 (0.127)
Black	−0.271 (0.151)	−0.270 (0.151)
Poverty	−0.077 (0.046)	−0.078 (0.046)
Female	0.173 (0.022)*	0.173 (0.022)*
English language learner	−0.113 (0.029)*	−0.113 (0.029)*
Special education	0.075 (0.040)	0.075 (0.040)
Gifted	0.018 (0.040)	0.018 (0.040)
Honor English	0.169 (0.056)*	0.169 (0.056)*
Baseline Peer ELA Z score	0.109 (0.057)	0.109 (0.057)
Baseline year mathematics Z score	0.189 (0.020)*	0.189 (0.020)*
Baseline year ELA Z score	0.439 (0.021)*	0.439 (0.021)*

Based on the dosage-dependent model, the average treated student received a 0.485 treatment dosage. Because of this, using the dosage-dependent model we could estimate an average treatment on the treated (ATT) at $(0.485 * 0.118) = 0.057$

* $p = .05$

Table 3 Cohort 2 LDC elementary school teacher effect estimates on 2017–2018 smarter balanced ELA performance, dosage-independent model

Variables	Dosage independent model coefficient (SE)
<i>Level 1—LDC teacher treatment</i>	0.026 (0.034)
<i>Level 2—Student characteristics</i>	
Hispanic	− 0.058 (0.028)*
Black	− 0.166 (0.040)*
Poverty	− 0.029 (0.019)
Female	0.100 (0.017)*
English Language Learner	− 0.110 (0.030)*
Special education	− 0.093 (0.031)*
Gifted	0.014 (0.034)
% teachers with < 3 years of experience	− 0.001 (0.000)
Baseline peer ELA Z score	0.056 (0.026)*
Baseline year mathematics Z score	0.223 (0.017)*
Baseline year ELA Z score	0.501 (0.017)*

*Significance at $p = .05$ **Table 4** Cohort 2 LDC middle school teacher effect estimates on 2017–2018 smarter balanced ELA performance

Variables	Dosage-dependent model coefficient (SE)	Dosage-independent model coefficient (SE)
<i>Level 1 LDC teacher treatment</i>	0.149 (0.063)*	0.055 (0.038)
<i>Level 2 Student characteristics</i>		
Hispanic	− 0.042 (0.034)	− 0.042 (0.034)
Black	− 0.178 (0.050)*	− 0.180 (0.050)*
Poverty	− 0.019 (0.014)	− 0.019 (0.014)
Female	0.104 (0.013)*	0.104 (0.013)*
English Language Learner	− 0.113 (0.025)*	− 0.112 (0.025)*
Special education	− 0.108 (0.028)*	− 0.108 (0.028)*
Gifted	0.059 (0.021)*	0.059 (0.021)*
Honors Course	0.146 (0.028)*	0.148 (0.028)*
% teachers with < 3 years of experience	− 0.000 (0.001)	− 0.000 (0.001)
Baseline peer ELA Z score	0.003 (0.028)	0.002 (0.028)
Baseline year mathematics Z score	0.220 (0.013)	0.221 (0.013)*
Baseline year ELA Z score	0.513 (0.013)*	0.513 (0.013)*

For the dosage dependent model, since the average treatment student received a 0.440 treatment dosage we could estimate an average treatment on the treated (ATET) at $0.44 * 0.149 = 0.066$

* $p = .05$

honors English courses performed at higher levels than did their peers taking standard English courses.

Since the great majority (99%) of elementary students taught by the Cohort 2 elementary teachers were connected with only a single teacher and findings for the

two models become nearly identical, we only present the dosage independent model in Table 3. As shown, no statistically discernible LDC effect on the student outcome was found. In other words, students taught by LDC elementary teachers scored similarly on the ELA test to their matched peers in the comparison group.

The effects of the covariates on student performance were similar to those reported in Table 2, but included some additional significant variables. In addition to baseline ELA and mathematics performance, English language learner status, and gender, the variables of Black ethnicity, Hispanic ethnicity, special education, and average baseline ELA achievement of a student's peers were all significant predictors of ELA performance and were in the expected directions.

In Table 4, we present Cohort 2 middle school results of models that are both dosage dependent and dosage independent. Dosage dependent model results indicate a statistically significant and positive LDC effect on the student outcome. Across all treatment students and ignoring their dosage of exposure to LDC teachers, they scored about 0.055 standard deviation higher than the comparison students but the difference was not statistically significant. In contrast, the dosage independent model did not yield a statistically significant LDC effect. Treatment students with exposure to LDC in all three core subjects were estimated to perform 0.149 standard deviations above matched comparison students and the difference was statistically significant. This result difference pointed to the importance of including dosage information in the analysis model as the treatment dosage was not universal across students.

The significant dosage dependent effect suggests that treatment students who were exposed to a greater amount of LDC instruction (via multiple participating teachers in different content areas) benefited more from the program than the treatment students with less exposure to LDC teachers. The full LDC intervention effect size was estimated to be $d=0.149$ for those treatment students taught by LDC teachers in all their three core content classes in both semesters in the school year. They on average scored 0.149 standard deviations higher in ELA test than their academically and demographically similar comparison schools at academically and demographically similar comparison schools. The effect was smaller than the treatment students with less exposure to LDC teachers. With the average exposure to LDC teachers to be 49%, the average LDC treatment effect in our sample was calculated to be 0.066 standard deviations.

We also summarize the quasi-experimental results and provide a lens by which the reader can contextualize the magnitude of the results in Figs. 1 and 2. The Figures present the dosage-dependent LDC teacher treatment effects and the 95% confidence interval using the effects and associated standard errors for each of the three analyses. Figure 1 depicts the estimated effects of LDC in the three samples on students exposed to LDC teachers in all three major content areas: ELA, social studies/history, and science. The effect sizes for these estimates can be best understood as the estimated effect of LDC under ideal conditions. Figure 2 depicts the estimated effect of LDC in the three samples on the average observed student, who in the middle school context had considerably less exposure to LDC teachers in her core content classes.

As can be seen in the Figures, the lower bound of the confidence intervals around the estimates for the effect of Cohort 2 middle school teachers is above zero. For the

Fig. 1 Treatment effect on 2017–2018 smarter balanced ELA scores with 95% confidence interval for students with full LDC dosage, by cohort

Fig. 2 Treatment effect on 2017–2018 smarter balanced ELA scores with 95% confidence interval for students with average LDC dosage, by cohort

other two analyses, the confidence intervals cross the zero line, and therefore the estimates are not statistically significant at the 95 percent level. Note that the confidence intervals are much wider for Cohort 1 middle school teachers, likely due to a considerably smaller sample size of teachers and associated students. This smaller sample size (caused in part by attrition of schools and teachers from the program) was certainly a factor in the precision of the estimate for Cohort 1 teachers.

To further help the reader contextualize the statistically significant effects for Cohort 2 middle school teachers, we utilize an approach developed by Hill et al. (2008), which involves benchmarking against average student gains over the course of a school year. The authors reviewed annual achievement gains in seven nationally normed reading assessments: CAT5, SAT9, Terra Nova-CTBS, Gates-MacGinitie, MAT8, Terra Nova-CAT, and SAT10. They found that students gained an average of 0.32 standard deviations from grade 5 to 6, 0.23 standard deviations from grade 6 to 7, and 0.26 standard deviations from grade 7 to 8. A simple mean of these three average gains is 0.27.

Using this benchmark, and assuming a 9-month school year, the 0.066 effect estimate for students with average observed LDC dosage is of a similar magnitude to 2.2 months of learning in the Hill et al. (2008) meta-analysis $[(0.066/0.27) * 9 = 2.2]$. Likewise, the 0.149 effect estimate for students with full LDC dosage aligns to approximately 5 months of schooling $[0.149/0.27 * 9 = 5]$. It is important to note again that the ideal conditions of students being exposed to LDC in all three core content areas across the whole school year was not met for most students; therefore, the extrapolation of 2.2 months is the figure best aligned with the actual observed effect of LDC.

Conclusions and Significance

The current paper reports on early student academic outcome results from a multi-year mixed methods study of the implementation and impact of LDC using a quasi-experimental design (QED), as implemented in one large urban school district. 2017–2018 was LDC's second year of implementation in the school district. Among the 45 schools implementing LDC in 2017–2018, 14 are Cohort 1 schools which began implementation during 2016–2017, and 31 are Cohort 2 schools, which commenced at the beginning of the 2017–2018 school year. LDC experienced a high level of attrition among Cohort 1 schools and participants; nearly one-third of schools dropped out of the program after 2016–2017, and within the remaining schools, nearly half of teachers did not continue with LDC in 2017–2018.

Sufficient sample size was available to conduct quasi-experimental tests of the effect of LDC on students under Cohort 1 middle school teachers with 2 years of LDC, and under Cohort 2 elementary and middle school teachers who started their first year of LDC. All effect estimates were in the positive direction, with a statistically significant effect found for Cohort 2 middle school teachers with 1 year of LDC when LDC treatment was treated as a continuous variable accounting for students' different levels of exposure to LDC teachers.

With Cohort 1 schools only retaining about one third of their original group of classroom teachers, it is no surprise that we didn't see any LDC effects for the Cohort 1 with 2 years of implementation. We are surprised to find a statistically significant LDC effect on student learning among Cohort 2 middle school teachers as the LDC model only expects a student effect from teachers having 2 years of experience implementing LDC in their classrooms. With that said, the results also indicate

that this significant effect could only be achieved if the team of teachers in ELA, science, and social science all participated together in the LDC implementation.

While the findings are encouraging, it is important to note that these are early results. The LDC team has been making continuous improvements and refinements to their program in the past few years to increase its effectiveness and by extension their participant retention. We are looking forward to conducting the final analysis based on the 2018–2019 student data and reporting on the LDC effect on student learning for both Cohorts 1 and 2 schools and teachers.

Acknowledgements Funding was provided by U.S. Department of Education (Grant No. 20145515).

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Ball, D. L. (2000). Bridging practices intertwining content and pedagogy in teaching and learning to teach. *Journal of Teacher Education*, 51(3), 241–247.
- Browne, W. J., Goldstein, H., & Rasbash, J. (2001). Multiple membership multiple classification (MMMC) models. *Statistical Modelling*, 1(2), 103–124.
- Carter, M., Ferzli, M., & Wiebe, E. N. (2007). Writing to learn by learning to write in the disciplines. *Journal of Business and Technical Communication*, 21(3), 278–302.
- Chadwick, J. A. (2015). *Common core: Paradigmatic shifts*. Cambridge: Cambridge Scholars Publishing.
- Darling-Hammond, L., Wilhoit, G., & Pittenger, L. (2014). Accountability for college and career readiness: Developing a new paradigm. *Education Policy Analysis Archives*. <https://doi.org/10.14507/epaa.v22n86.2014>
- De La Paz, S., Monte-Sano, C., Felton, M., Croninger, R., Jackson, C., & Piantedosi, K. W. (2016). A historical writing apprenticeship for adolescents: Integrating disciplinary learning with cognitive strategies. *Reading Research Quarterly*, 52(1), 31–52.
- Drapper, R. J. O. (2008). Redefining content-area literacy teacher education: Finding my voice through collaboration. *Harvard Educational Review*, 78(1), 60–83.
- Duesbery, L., Werblow, J., & Twyman, T. (2011). *The effect of the Seeds of Science/Roots of Reading curriculum (planets and moons unit) for developing literacy through science in fifth grade*. Berkeley, CA: university of California Berkeley, Graduate School of Education. Retrieved from http://learningdesigngroup.org/sites/default/files/resources/files/final_planets_and_moons.pdf. Accessed 4 May 2021.
- Fix, E., & Hodges, J. (1951). *Discriminatory analysis, nonparametric discrimination: Consistency properties (Technical Report 4)*. Randolph Field, TX: USAF School of Aviation Medicine.
- Goldschmidt, M. (2015). Teaching writing in the disciplines: Student perspectives on learning genre. *Teaching & Learning Inquiry*, 2(2), 25–40.
- Graff, N. (2011). “An effective and agonizing way to learn”: Backwards design and new teachers’ preparation for planning curriculum. *Teacher Education Quarterly*, 38(3), 151–168.
- Grossman, P., Hammerness, K., & McDonald, M. (2009). Redefining teaching, re-imagining teacher education. *Teachers and Teaching*, 15(2), 273–289.

- Herman, J., Epstein, S., & Leon, S. (2016). Supporting Common Core instruction with literacy design collaborative: A tale of two studies. *AERA Open*, 2(3), 2332858416655782.
- Hill, C. J., Bloom, H. S., Black, A. R., & Lipsey, M. W. (2008). Empirical benchmarks for interpreting effect sizes in research. *Child Development Perspectives*, 2(3), 172–177.
- Iacus, S. M., King, G., & Porro, G. (2011). Causal inference without balance checking: Coarsened exact matching. *Political Analysis*, 20(1), 1–24. <https://doi.org/10.1093/pan/mpi013>
- Klein, P. D., & Kirkpatrick, L. C. (2010). A framework for content area writing: Mediators and moderators. *Journal of Writing Research*, 2(1), 1–46.
- Levin, S., Leow, C., & Poglinco, S. M. (2013). Scale-up and sustainability study of the LDC and MDC initiatives. In *Research for action*.
- Monte-Sano, C., & De La Paz, S. (2012). Using writing tasks to elicit adolescents' historical reasoning. *Journal of Literacy Research: JLR*, 44(3), 273–299.
- Reisman, A. (2012). The “Document-Based Lesson”: Bringing disciplinary inquiry into high school history classrooms with adolescent struggling readers. *Journal of Curriculum Studies*, 44(2), 233–264.
- Shanahan, T., & Shanahan, C. (2008). Teaching disciplinary literacy to adolescents: rethinking content-area literacy. *Harvard Educational Review*, 78(1), 40–59.
- Tranmer, M., Steel, D., & Browne, W. J. (2014). Multiple-membership multiple-classification models for social network and group dependences. *Journal of the Royal Statistical Society Series A (Statistics in Society)*, 177(2), 439–455. <https://doi.org/10.1111/rssa.12021>
- Wang, J., Neskovic, P., & Cooper, L. N. (2007). Improving nearest neighbor rule with a simple adaptive distance measure. *Pattern Recognition Letters*, 28(2), 207–213.
- Wiggins, G. P., & McTighe, J. (2005). *Understanding by design*. ASCD.
- Zhu, P., Jacob, R., Bloom, H., & Xu, Z. (2012). Designing and analyzing studies that randomize schools to estimate intervention effects on student academic outcomes without classroom-level information. *Educational Evaluation and Policy Analysis*, 34(1), 45–68.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.